B

~~Description~~

A Vector for Gene Trap, and A Method for Gene Trapping by Using The Vector

Ins a'
Ins B'

~~Technical Field~~

The present invention relates to a new vector system to facilitate the cloning and functional analysis of new genes of a fly, *Drosophila melanogaster*, and a method for gene trapping with the vector system.

Ins B2

~~Background Art~~

There are numerous examples for application of gene trapping methods in wide range of living organisms including maize and mouse (Gossler et al,. Science, 244:463-465, 1989).

With respect to tools for gene trapping, the application of different types of enhancer trap P-element vectors (Wilson et al., Genes & Development, 3:1301-1313, 1989) for cloning and analyzing trapped genes, as well their use for mosaic analysis with the help of the Gal4/UAS transcription activator system has proven fruitful. However, sometimes the expression pattern of the Gal4 or other reporter gene of the vector construct is affected by enhancers belonging to more than one gene. Similarly, in some cases it is difficult to determine whether the enhancer trap insertion effects the function of one or more of the neighboring genes.

These circumstances altogether with the fact that in some cases the mutant phenotype could be attributed to the

changed expression of a gene with its nearest exon located more than 30 kB apart from the insertion site, can lead in unfortunate cases to an ordeal when it's time to clone and analyze the affected gene.

One object of this application is to provide a vector system that includes specifically designed artificial regulatory sequences as well as selection methods for easy screening of positive recombinant lines. More especially, this application intends to provide a vector system of this invention offering much easier and faster cloning opportunities of the affected gene, compared to the widely used enhancer trap P-element vectors. Another object of this application is to provide easier detection method possibilities of the successful trapping events and much higher chance to get more characteristic ("functional") expression patterns of the reporter gene because in the contrary with much of the cases with enhancer trap lines, when using the vector system of this invention, the reporter gene expression is influenced only by a single endogenous transcription unit and effects only the expression of the very same gene.

Summary ~~Disclosure~~ of Invention

The first invention of this application is a vector for trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster;* and

a synthetic splicing donor site.

One embodiment of the first invention is that the recombinant plasmid is derived from pCasper3.

Other embodiments of the first invention are that the reporter gene is the Gal4 gene, Gal4 DNA binding domain-P53 fusion gene or the Gal4-firefly luciferase fusion gene.

Further embodiment of this first invention is that the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene.

More further embodiment of the first invention is that the drug resistance gene is neomycin-phosphotranspherase gene and its promoter is a heatshock promoter.

The second invention of this application is a method for trapping an unknown gene of *Drosophila melanogaster* by using a vector which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster;* and

a synthetic splicing donor site,

which method comprises the steps of:

(a) introducing the vector into the genome of a white minus fly;

(b) selecting primary transformants resistant to a drug;

(c) crossing the primary transformants with a transposase source strain to force the vector to jump into other locations;

(d) selecting secondary transformants by picking up the flies having strong eye color,

(e) crossing the secondary transformants with UAS (Upstream Activator Sequence)-luciferase harboring strain and measuring the reporter gene expression of the resultant flies; and

(f) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

The third invention of this application is a method for trapping an unknown gene of *Drosophila melanogaster* by using a vector A which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

Gal4 DNA binding domain-P53 fusion gene as a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and

a synthetic splicing donor site,

and a vector B derived from pCasperhs, which has the heatshock promoter directed Gal4 activator domain-large T antigen fusion gene within polycloning site of the pCasperhs,

which method comprises the steps of:

(a) introducing each of the vectors A and B into the genomes of separate white minus flies;

(b) selecting primary transformants for the vector A which are resistant to the drug, and selecting primary transformants for the vector B which have an eye color;

(c) crossing the primary transformants for the vector A with a transposase source strain to force the vector to jump into other locations;

(d) selecting secondary transformants for the vector A by picking up the flies having strong eye color;

(e) crossing the secondary transformants with the primary transformants for the vector B to obtain flies harboring both the vectors A and B;

(f) crossing the flies obtained in the step (e) with an UAS-luciferase harboring fly strain and measuring the reporter gene expression of the resultant flies after a heatshock treatment; and

(g) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

Embodiments of the second and third inventions are corresponded to the embodiments of the first invention, and they will be more precisely described in the following description.

Brief Description of The Drawings

Figure 1 shows the schematic map of the vector of this invention, pTrap-hsneo.

Figure 2 shows the schematic map of the vector of this invention, pTrap-G4-p53.

Figure 3 shows the schematic map of the vector of this

invention, pCasperhs-G4-LT.

Figure 4 shows the schematic map of the vector of this invention, pTrap-G4-luc.

Sub B3 Figure 5 shows the shematic drawing of a fly genome to which the vector of this invention is inserted for cloning.

Figure 6 shows the results of sequencing RT-PCR products of aop-Gal4 and m-white-aop fusion mRNAs.

Figure 7 presents pictures of characteristic beta-galactosidase staining patterns in different parts of the fly brain resulted from crossing positive gene trap lines with flies harboring a UAS-lacZ construct.

B

Description of the Preferred Embodiments ~~Best Mode for Carrying Out the Invention~~

A vector construct of the first invention, for example, can be based on the commonly used, P-element transformation vector, pCasper3 (Pirotta, Vectors: A survey of molecular cloning vectors and their uses, eds. Rodriguez, R.L. & Denhardt, D.T., Butterworths, Boston. 437-456, 1998) and the convenient Gal4-UAS expression system (Brand and Perrimon, Development, 118:401-415, 1993).

A promoterless Gal4 gene preceded by an artificial consensus splicing acceptor site and a synthetic "stop/start" sequence to govern the read through translation coming from upstream exon(s) of the trapped gene into the proper reading frame of Gal4 was inserted into the polycloning site of pCasper3.

The removal of the whole 3' UTR (untranslated region) sequence of the mini-white gene and replacement by an artificial splicing donor site resulted in a truncated gene without its own poly-adenylation site.

Without a successful gene trapping event this truncated mini-white gene was not expected to confer any eye color, therefore in this invention a heatshock promoter directed neomycin-phosphotransferase (hs-neo) gene for helping selection of primary transformants by antibiotic feeding has been inserted.

Figure 1 shows the schematic map of the gene trap construct (pTrap-hsneo), and SEQ ID No.1 is the complete nucleotide sequence of the vector pTrap-hsneo.

Another gene trap construct, pTrap-G4-p53 (Figure 2) is created by replacing the Gal4 coding sequence of plasmid pTrap-hsneo with a Gal4 DNA binding domain-P53 fusion gene (Clontech, Matchmaker Two Hybrid System, #K1605-1). When this construct coexists in the genome of the same fly with another vector, pCasperhs-G4-LT (Figure 3) containing a heatshock promoter directed Gal4 activator domain-large T antigen (Clontech, Matchmaker Two Hybrid System, #K1605-1) fusion gene, the assembly of a functional Gal4 molecule, through p53-large T antigen interaction, can be regulated by external heatshock.

Sub B4 On this way, the possibility of an intentional temporary control of Gal4 activity became available. In other words the Gal4 expression in a pattern as already determined spatially by the promoter of the trapped gene now can be induced at any desired stage of development by external heatshock.

In order to make the detection of Gal4 expression easier, the Gal4 gene in another construct is replaced with a Gal4-firefly luciferase fusion gene to get pTrap-G4-luc (Figure 4). This artificial gene is coding for a fusion

polypeptide which has preserved both enzymatic activities.

The easy measuring of luciferase activity by luminoassay (Brandes et al., Neuron, 16:687-694, 1996) makes the detection of Gal4 activity comfortable in every single living fly.

Then, one of the best mode of the second or third invention, a method for gene trapping using the vector system, is described in detail.

(1) Screening:

The gene trap vector constructs can be introduced into the genome of a white minus fly by microinjection. The selection of primary transformants is possible by using G418, an analog of neomycin, resistance conferred by hs-neo gene. (When performing transformation experiments with these constructs it's turned out that the truncated mini-white gene generally provides a very slight yellow eye color which could be distinguished from w-minus phenotype in most of the cases, therefore G418 selection apparently is not necessary.)

After a line with the gene trap construct is being established, the secondary transformants can be generated on the usual way by crossing the original line with a so-called jumpstarter containing the transposase expressing delta 2-3 genetic element.

Usually a certain percentage, between 4 and 8, of the secondary transformants have much stronger eye color (deep orange or reddish) than the ancestor fly indicating that the construct was being inserted downstream of a promoter and now the mini-white gene is using the transcriptional "facilities" of that gene (e.g.: poly-adenylation site and transcriptional

terminator) instead of its removed ones. They are the most likely candidates for successful gene trap events. In case of these lines the vector probably has been inserted either into an intron of a gene or upstream from the first intron into the 5′ UTR in proper orientation (that is the direction of transcription is same for the "trapped gene" and the mini-white (and Gal4) genes as well). The mini-white gene has its own promoter therefore its expression pattern is supposed to be largely independent from that of the trapped gene.

These positive lines are to be checked in the next step for Gal4 expression by crossing them with a "marker" line harboring a UAS-luciferase reporter gene construct. (When using pTrap-G4-luc vector, this step is obviously not necessary.) Usually very strong correlation was found between eye color and Gal4 expression: more than 90% of the lines having strong eye color proved to be expressing Gal4 by means of luciferase assay using luminometer (Brandes et al., Neuron, 16:687-692, 1996).

(2) Cloning:

Sub B5 When the gene trap construct is being inserted into an intron of an endogenous gene, the marker genes of the construct are supposed to be spliced on mRNA level to the exons of the trapped gene by using the artificial splicing acceptor and donor sites. More exactly while the Gal4 mRNA should be joint to the exon(s) located upstream of the insertion site, at the same time the mini-white mRNA is fused to the following exon(s) accomplishing the dual tagging of the trapped gene (Figure 5).

This feature can be used for quickly and easily

identifying the trapped gene by means of 3′ and 5′ RACE (Rapid Amplification of cDNA Ends) experiments. Even cloning and sequencing only a part of the caught mRNA still provides reasonable chance to find homologous mRNAs in the BDGP (Berkeley Drosophila Genome Project) EST (Expressed Sequence Tag) library.

With these approaches, the identification of an already cloned gene can take less then a week compared to the usually more than one year period in average when analyzing a mutant created by some enhancer trap construct.

It's well-known from the literature and the present inventors also have experienced that P-element vectors tend to integrate into or near the 5′ UTR of active genes. (The present inventors found that in these cases if the insertion occurred upstream from the first intron, and therefore the artificial splicing acceptor site could not be utilized, the Gal4 gene was expressed by read-through transcription from the nearby promoter.)

The advantage of this tendency can be taken by cloning and sequencing the flanking genomic sequences of the insertion site by inverse or vectorette PCR or by plasmid rescue using suitable restriction digestion to recover the neomycin resistance gene of the construct. Then again the BDGP library can be searched to find any significant matching.

(3) Rescue:

The only reliable way to confirm that any observed mutant phenotype is really the consequence of the P-element insertion is to rescue that particular phenotype. Expectedly the phenotype (some alteration from wild type fly)is caused

by changed expression of gene(s) disturbed by insertion of the P-element. The rescue can be made by expressing the cDNA of the suspected gene most preferable with identical spatial and temporary pattern than that of the gene itself.

As it was expected, the vector constructs of the first invention usually cause strong phenotypes. It's not surprising at all because the trapped genes are supposed to be split into two parts on mRNA level resulting in null mutants in majority of the cases. Accordingly mutants obtained by this method frequently show homozygous lethality or sterility. Hypomorphic mutants can be obtained by forcing imprecise excision of the gene trap P-element construct.

Sub B6 As mentioned above, the Gal4 expression is obliged to reflect precisely to that of the trapped gene simply because the Gal4 gene has no its own promoter and they share a common, fused mRNA.

This identical expression provides unique opportunity to rescue the mutant phenotype by crossing this fly with another one harboring the UAS directed, cloned cDNA of the trapped gene.

Sub B7 On this way either the original, homozygous null mutant gene trap fly or any transheterozygous derivative of that with some hypemorphic allele over the null mutant allele can be rescued.

(4) Determination of spatial and developmental expression pattern of the trapped gene:

Histochemical determination of the spatially and temporarily controlled expression of any trapped gene is also easy following introduction of a UAS-lacZ construct into the

genome of the same fly and performing either X-gal or antibody staining for beta-galactosidase.

(5) Mosaic analysis:

Possession of a large collection of fly lines with different, characteristic and, in the case of the pTrap-G4-p53/pCasperhs-G4-TL vector system, inducible Gal4 expression pattern makes feasible carrying out mosaic analysis of virtually any gene of interest by directing the expression of their UAS-constructs on a mutant background with different Gal4 expression patterns.

This approach can answer the question of where and when that particular gene is required to be expressed to rescue the mutant phenotype.

Similarly, any gene can be expressed in different ectopic patterns to generate new dominant mutant phenotypes. This approach might help to conclude the role of that particular gene and to identify the pathway, in which it's involved.

Example

The following example illustrates a specific embodiment of the various aspects of the invention. This example is not intended to limit the invention in any manner.

Figure 6 shows the results of sequencing RT-PCR products of aop-Gal4 and m-white-aop fusion mRNAs.

The template was total RNA prepared from a positive gene trap line which has the vector pTrap-hsneo being integrated into the first intron of the well-known aop (anterior open/pokkuri/yan) developmental gene. The sequences confirm that both splicing occurred precisely at that

particular nucleotides of the artificial regulatory sequences where it was expected.

On Figure 7, there are pictures of characteristic beta-galactosidase staining patterns in different parts of the fly brain resulted from crossing positive gene trap lines with flies harboring a UAS-lacZ construct.

~~Industrial Applicability~~

The vector system of this invention offers an exceptional opportunity for easy and fast cloning of the gene responsible for the observed phenotype. Furthermore, by using the UAS-driven coding sequence of any gene of interest, that particular gene can be expressed in identical patterns than those of the trapped genes and these expressions can be regulated temporarily at any desired developmental stage.

Sequence Listing

<110> Japan Science and Technology Corporation

<120> A Vector for Gene Trap, and A Method for Gene Trapping by Using The Vector

<150> Japan, Application No. 10-141952

<151> 22 May 1998

<160> 1

<170> PatentIn Ver. 2.0

<210> 1

<211> 11206

<212> DNA

<213> Artificial sequence

<220>

<221> 3'P sequence